

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/524,827A
Source: IFW16
Date Processed by STIC: 4/20/07

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 04/20/2007

PATENT APPLICATION: US/10/524,827A

TIME: 14:29:22

Input Set : A:\Revised sequence list-13173-00006-US.txt

Output Set: N:\CRF4\04202007\J524827A.raw

```

4 <110> APPLICANT: Sauer, Matt
5      Flachmann, Ralf
6      Klebsattel, Martin
7      Schopfer, Christel R
9 <120> TITLE OF INVENTION: Process for preparing ketocarotenoids in genetically
modified
10      organisms
12 <130> FILE REFERENCE: 13173-00006-US
14 <140> CURRENT APPLICATION NUMBER: US 10/524,827A
15 <141> CURRENT FILING DATE: 2005-02-18
17 <150> PRIOR APPLICATION NUMBER: PCT/EP2003/009106
18 <151> PRIOR FILING DATE: 2003-08-18
20 <150> PRIOR APPLICATION NUMBER: DE 102 38 980.2
21 <151> PRIOR FILING DATE: 2002-08-20
23 <150> PRIOR APPLICATION NUMBER: DE 102 38 978.0
24 <151> PRIOR FILING DATE: 2002-08-20
26 <150> PRIOR APPLICATION NUMBER: DE 102 38 979.9
27 <151> PRIOR FILING DATE: 2002-08-20
29 <150> PRIOR APPLICATION NUMBER: DE 102 53 112.9
30 <151> PRIOR FILING DATE: 2002-11-13
32 <150> PRIOR APPLICATION NUMBER: DE 102 58 971.2
33 <151> PRIOR FILING DATE: 2002-12-16
35 <160> NUMBER OF SEQ ID NOS: 82
37 <170> SOFTWARE: PatentIn version 3.4
40 <210> SEQ ID NO: 1
42 <211> LENGTH: 777
44 <212> TYPE: DNA
46 <213> ORGANISM: Nostoc sp. Strain PCC7120
48 <220> FEATURE:
50 <221> NAME/KEY: CDS
52 <400> SEQUENCE: 1
53 atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta      48
54 Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
55 1          5          10          15
57 ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt      96
58 Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe
59          20          25          30
61 att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta      144
62 Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
63          35          40          45
65 ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc      192
66 Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala
67          50          55          60
69 atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat      240

```

RAW SEQUENCE LISTING

DATE: 04/20/2007

PATENT APPLICATION: US/10/524,827A

TIME: 14:29:22

Input Set : A:\Revised sequence list-13173-00006-US.txt

Output Set: N:\CRF4\04202007\J524827A.raw

```

70 Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
71 65              70              75              80
73 gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat      288
74 Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn
75              85              90              95
77 ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa      336
78 Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys
79              100              105              110
81 gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat      384
82 Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp
83              115              120              125
85 tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg      432
86 Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp
87              130              135              140
89 tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga      480
90 Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly
91 145              150              155              160
93 tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa      528
94 Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Gln
95              165              170              175
97 aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta      576
98 Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
99              180              185              190
101 caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt      624
102 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
103              195              200              205
105 ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt      672
106 Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
107              210              215              220
109 tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac      720
110 Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
111 225              230              235              240
113 gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata      768
114 Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
115              245              250              255
117 tct tta taa      777
118 Ser Leu
122 <210> SEQ ID NO: 2
124 <211> LENGTH: 258
126 <212> TYPE: PRT
128 <213> ORGANISM: Nostoc sp. Strain PCC7120
130 <400> SEQUENCE: 2
131 Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
132 1              5              10              15
135 Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe
136              20              25              30
140 Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
141              35              40              45
144 Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala

```

DATE: 04/20/2007

TIME: 14:29:22

Input Set : A:\Revised sequence list-13173-00006-US.txt

Output Set: N:\CRF4\04202007\J524827A.raw

	50					55					60						
145	Met	Leu	Trp	Gln	Thr	Phe	Leu	Tyr	Thr	Gly	Leu	Phe	Ile	Thr	Ala	His	
148	65					70					75					80	
152	Asp	Ala	Met	His	Gly	Val	Val	Tyr	Pro	Lys	Asn	Pro	Arg	Ile	Asn	Asn	
153					85					90					95		
156	Phe	Ile	Gly	Lys	Leu	Thr	Leu	Ile	Leu	Tyr	Gly	Leu	Leu	Pro	Tyr	Lys	
157				100					105					110			
160	Asp	Leu	Leu	Lys	Lys	His	Trp	Leu	His	His	Gly	His	Pro	Gly	Thr	Asp	
161				115				120					125				
164	Leu	Asp	Pro	Asp	Tyr	Tyr	Asn	Gly	His	Pro	Gln	Asn	Phe	Phe	Leu	Trp	
165		130					135					140					
168	Tyr	Leu	His	Phe	Met	Lys	Ser	Tyr	Trp	Arg	Trp	Thr	Gln	Ile	Phe	Gly	
169	145					150					155					160	
172	Leu	Val	Met	Ile	Phe	His	Gly	Leu	Lys	Asn	Leu	Val	His	Ile	Pro	Glu	
173				165						170					175		
176	Asn	Asn	Leu	Ile	Ile	Phe	Trp	Met	Ile	Pro	Ser	Ile	Leu	Ser	Ser	Val	
177				180					185					190			
180	Gln	Leu	Phe	Tyr	Phe	Gly	Thr	Phe	Leu	Pro	His	Lys	Lys	Leu	Glu	Gly	
181		195						200					205				
184	Gly	Tyr	Thr	Asn	Pro	His	Cys	Ala	Arg	Ser	Ile	Pro	Leu	Pro	Leu	Phe	
185		210					215					220					
188	Trp	Ser	Phe	Val	Thr	Cys	Tyr	His	Phe	Gly	Tyr	His	Lys	Glu	His	His	
189	225					230					235				240		
192	Glu	Tyr	Pro	Gln	Leu	Pro	Trp	Trp	Lys	Leu	Pro	Glu	Ala	His	Lys	Ile	
193				245						250					255		
196	Ser	Leu															
200	<210> SEQ ID NO: 3																
202	<211> LENGTH: 789																
204	<212> TYPE: DNA																
206	<213> ORGANISM: Nostoc punctiforme																
208	<220> FEATURE:																
210	<221> NAME/KEY: CDS																
213	<400> SEQUENCE: 3																
214	ttg	aat	ttt	tgt	gat	aaa	cca	gtt	agc	tat	tat	gtt	gca	ata	gag	caa	48
215	Leu	Asn	Phe	Cys	Asp	Lys	Pro	Val	Ser	Tyr	Tyr						

RAW SEQUENCE LISTING

DATE: 04/20/2007

PATENT APPLICATION: US/10/524,827A

TIME: 14:29:22

Input Set : A:\Revised sequence list-13173-00006-US.txt

Output Set: N:\CRF4\04202007\J524827A.raw

```

235 Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
236      85      90      95
238 cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag      336
239 Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
240      100      105      110
242 aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat      384
243 Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
244      115      120      125
246 ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc      432
247 Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
248      130      135      140
250 atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta      480
251 Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
252 145      150      155      160
254 ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc      528
255 Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
256      165      170      175
258 tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat      576
259 Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
260      180      185      190
262 ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat      624
263 Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
264      195      200      205
266 ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc      672
267 Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
268      210      215      220
270 gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat      720
271 Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
272 225      230      235      240
274 gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac      768
275 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
276      245      250      255
278 aat tca gta acc aat tcg taa      789
279 Asn Ser Val Thr Asn Ser
280      260
283 <210> SEQ ID NO: 4
285 <211> LENGTH: 262
287 <212> TYPE: PRT
289 <213> ORGANISM: Nostoc punctiforme
291 <400> SEQUENCE: 4
292 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
293 1      5      10      15
296 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
297      20      25      30
300 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
301      35      40      45
304 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
305      50      55      60
308 Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His

```

RAW SEQUENCE LISTING

DATE: 04/20/2007

PATENT APPLICATION: US/10/524,827A

TIME: 14:29:22

Input Set : A:\Revised sequence list-13173-00006-US.txt

Output Set: N:\CRF4\04202007\J524827A.raw

```

309 65          70          75          80
312 Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
313          85          90          95
316 Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
317          100          105          110
320 Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
321          115          120          125
324 Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
325          130          135          140
328 Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
329 145          150          155          160
332 Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
333          165          170          175
336 Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
337          180          185          190
340 Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
341          195          200          205
344 Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
345          210          215          220
348 Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
349 225          230          235          240
352 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
353          245          250          255
356 Asn Ser Val Thr Asn Ser
357          260
360 <210> SEQ ID NO: 5
362 <211> LENGTH: 762
364 <212> TYPE: DNA
366 <213> ORGANISM: Nostoc punctiforme
368 <220> FEATURE:
370 <221> NAME/KEY: CDS
372 <400> SEQUENCE: 5
373 gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca      48
374 Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
375 1          5          10          15
377 gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc      96
378 Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
379          20          25          30
381 att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac      144
382 Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp
383          35          40          45
385 atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa      192
386 Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
387          50          55          60
389 aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat      240
390 Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
391 65          70          75          80
393 ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca      288
394 Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr

```

VERIFICATION SUMMARY

DATE: 04/20/2007

PATENT APPLICATION: US/10/524,827A

TIME: 14:29:23

Input Set : A:\Revised sequence list-13173-00006-US.txt

Output Set: N:\CRF4\04202007\J524827A.raw